

Trial Over Error

Enhancing lead prioritization in clinical trial services



Natalie Chuang



Tessie Xie



Thermo Fisher Advisor: Molly Isermann

Faculty Advisor: Stephen Graves

About Thermo Fisher Scientific



>\$40B
in annual revenue



8 brands
of products and services



57%
of clients are in pharma/biotech



Global
client and site presence

The Clinical Trials Division (CTD) within Thermo Fisher Scientific supports end-to-end management of pharma services for every stage of drug development.

The **clinical ancillary management (CAM)** business provides sourcing and supply solutions for consumables, equipment, or devices to support clinical trial operation.

Problem Statement

Since becoming a distinct service line in CTD, a **limited understanding of purchase behavior** has been recognized in the CAM business as it relates to clinical trial and customer features.



Identify factors that affect CAM sales...



...estimate CAM **business value**...



...and **prioritize** lead prospects.

Data



Salesforce Purchase Data

- Transaction records for Clinical Trials Division purchases
- From **2019 to present**



Customer Data

- Customer **purchase behavior**
- Customer **industry sector**



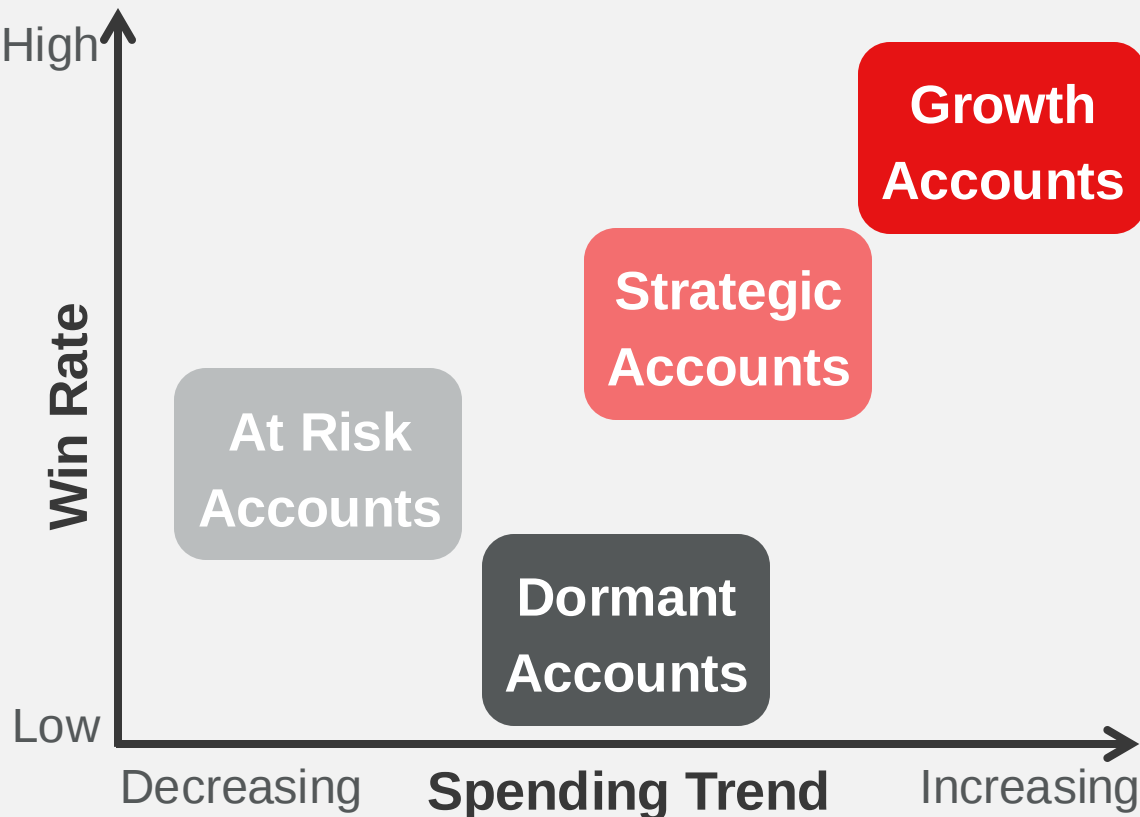
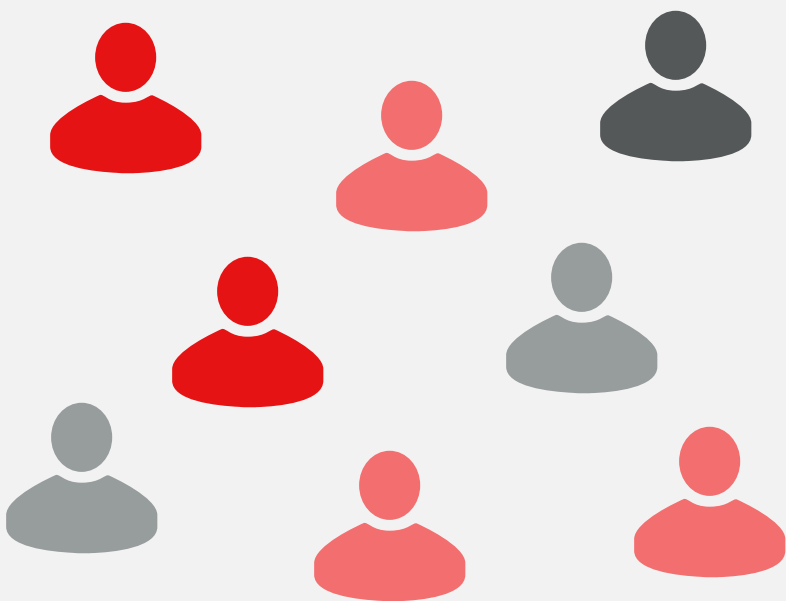
Clinical Trials Data

- Protocol and therapy features** from multiple registries globally
- Industry-sponsored** trials

Methodology

1. Customer Segmentation

Group customer accounts using K-means clustering by similar purchasing behavior



2. Prospect Identification

Use binary classification to screen out trials with low CAM potential

Model	AUC	Recall
K-Nearest Neighbors	0.575	0.202
Logistic Regression	0.621	0.807
CART	0.611	0.790
Random Forest	0.677	0.630
XGBoost	0.654	0.790

Model with highest recall chosen to avoid missing potential sale opportunities

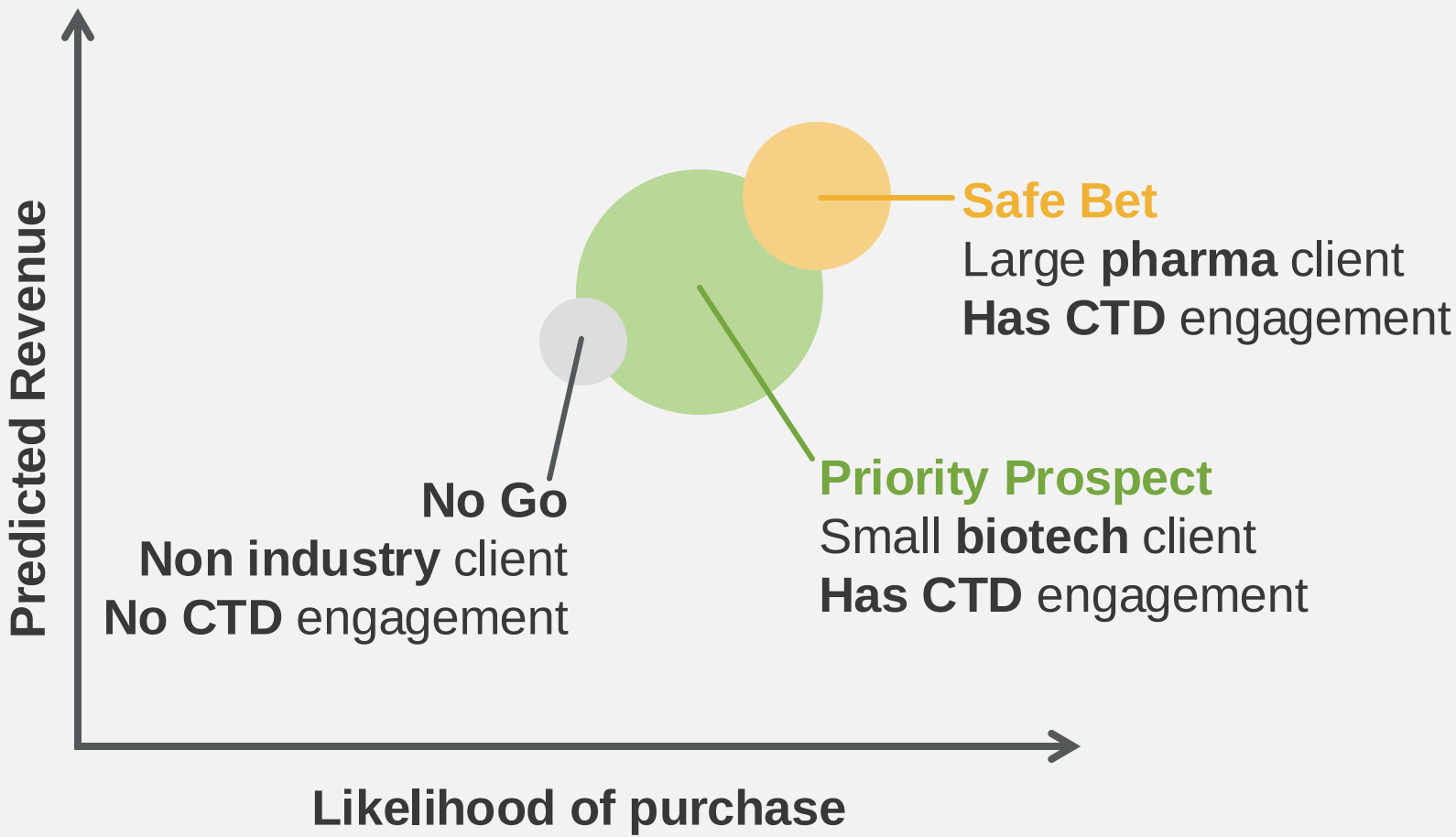
3. Prospect Scoring

Win Propensity
Quantify likelihood of winning a sale opportunity using binary classification

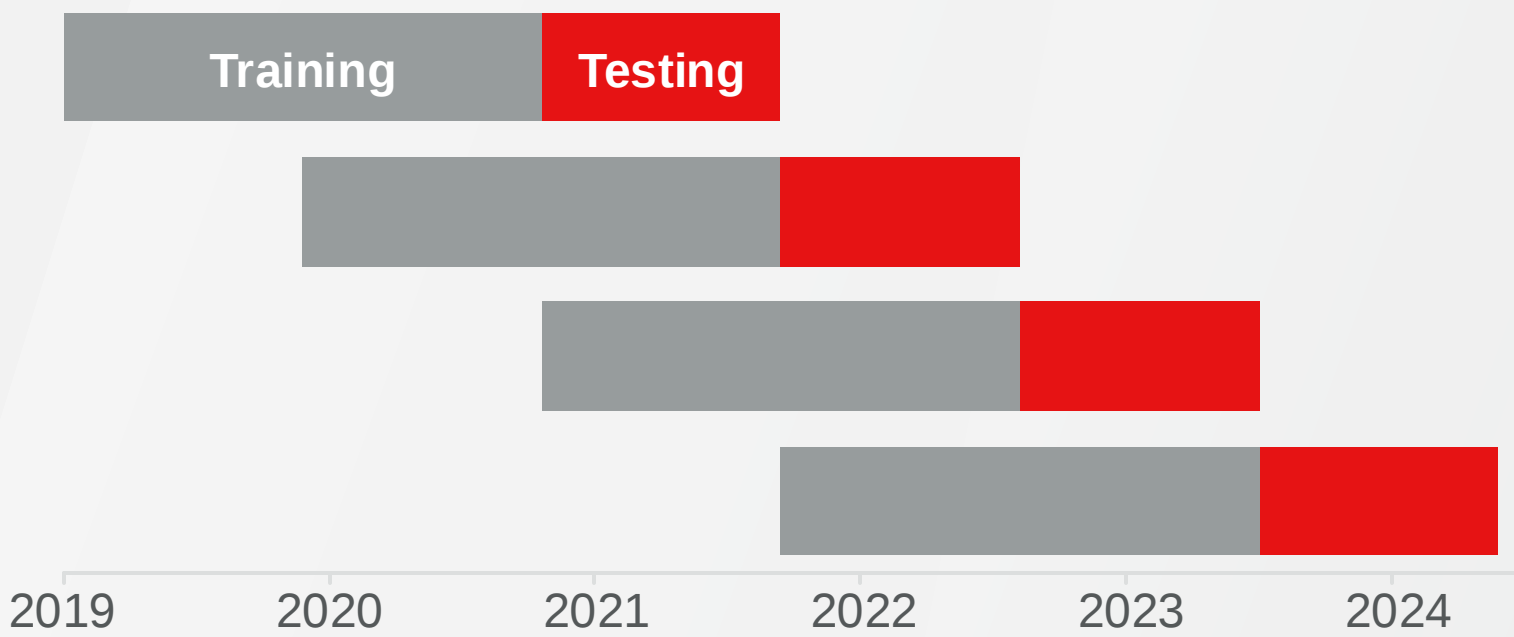
Revenue Prediction
Predict high, low, or zero revenue with multiclass classification

Feature Weighting
Give weights to target market features based on CAM business priorities

CAM Prioritization Score



Results & Validation



Model predictions for win propensity and revenue

vs

Baseline obtained from historic win rate and revenue distribution

23%

Increase in accuracy over baseline

97%

Of revenue-generating sales correctly identified

Top Revenue Drivers

- Infectious disease trials
- Injectable drugs
- Phase III trials
- Large pharma sponsors

Impact



Expanded sales funnel to grow CAM pipeline



Strategy to target key emerging markets



Framework to rank and prioritize prospects